AD-A086 409   MOORE SCHOOL OF ELECTRICAL ENGINEERING PHILADELPHIA PA   F/G 12/1
                ROBUST AND QUANTIZED WIENER FILTERS FOR P-POINT SPECTRAL CLASSE--ETC(U)
                1980   L J CIMINI, S A KASSAM   AFOSR-77-3154
UNCLASSIFIED   AFOSR-TR-80-0425   NL

AL A 086-409							END DATE FILMED 8-80 DTIC

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 80-0425	2. GOVT ACCESSION NO. AD-A086 409	3. RECIPIENT'S CATALOG NUMBER

4. TITLE *(and Subtitle)* ROBUST AND QUANTIZED WIENER FILTERS FOR P-POINT SPECTRAL CLASSES	5. TYPE OF REPORT & PERIOD COVERED Interim
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s) L. J. Cimini and S. A. Kassam	8. CONTRACT OR GRANT NUMBER(s) AFOSR 77-3154

9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Pennsylvania Department of Systems Eng. Philadelphia, PA 19104	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F

11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling Air Force Base, Washington D.C 20332	12. REPORT DATE 1980
	13. NUMBER OF PAGES 6

14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*	15. SECURITY CLASS. *(of this report)* UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

Conference on Information Sciences and Systems, Princeton University, March 1980; paper to appear in proceedings.

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Wiener Filters, Robust Filters, Spectral Density

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

The classical solution for linear estimation of a signal in additive noise is given by the well-known Wiener filter. In certain circumstances, however, it is advantageous to have a filter which is piecewise constant and which provides good performance over an entire class of signal and noise power spectral densities. First, a procedure is presented for determining the optimal piecewise constant filter. Then, for a particular class of power spectral densities, termed the p-point spectral class, the robust filter is determined.

DD FORM 1 JAN 73 1473

Princeton
Conf. 1980

ROBUST AND QUANTIZED WIENER FILTERS FOR p-POINT SPECTRAL CLASSES

Leonard J. Cimini
Saleem A. Kassam
University of Pennsylvania
The Moore School of Electrical Engineering
Philadelphia, PA 19104

Interim rept.,

AFOSR-77-315

ABSTRACT

The classical solution for linear estimation of a signal in additive noise is given by the well-known Wiener filter. In certain circumstances, however, it is advantageous to have a filter which is piecewise constant and which provides good performance over an entire class of signal and noise power spectral densities. First, a procedure is presented for determining the optimum piecewise constant filter. Then, for a particular class of power spectral densities, termed the p-point spectral class, the robust filter is determined.

I. INTRODUCTION

In the linear estimation of a random signal in additive noise, an estimate, $\hat{s}(t)$, of the signal, $s(t)$, is obtained by passing the received signal, $x(t) = s(t) + n(t)$, through a linear filter with impulse response, $h(t,\tau)$. Assuming the signal and noise processes are real, zero-mean, uncorrelated and wide-sense stationary with power spectral densities (PSD's), $S(w)$ and $N(w)$, respectively, the Wiener filter solution is obtained by choosing $h_o(t)$, the impulse response of the optimum filter, to minimize the mean-square-error (MSE) between the signal and its estimate. The Fourier transform, $H_o(w)$, of $h_o(t)$ is the solution to

$$e(S,N;H_o) = \min_{H} e(S,N;H) \qquad (1)$$

where the MSE, $e(S,N;H)$, is given by

$$e(S,N;H) = \frac{1}{2\pi}\int_{-\infty}^{\infty} [S(w)|1-H(w)|^2 + N(w)|H(w)|^2]dw \qquad (2)$$

This is for the case where $H_o(w)$ is not constrained to be causal.

The solution to (1) is found to be

$$H_o(w) = \frac{S(w)}{S(w)+N(w)} \qquad (3)$$

the well-known Wiener filter. Knowing the signal and noise PSD's, one can thus obtain the optimum Wiener filter.

Section II details a procedure for obtaining a piecewise constant filter which minimizes the MSE. By deriving the necessary conditions for minimization, the optimum piecewise constant (OPC) filter is shown to be distinct from the optimum quantization (OQ) of $H_o(w)$. The OPC filter provides a simplified filter structure with only a small degradation in performance.

In Section III, we show that a piecewise constant filter also possesses the property of robustness for a particular class of PSD's - the p-point spectral class. A proof of the robustness of the piecewise constant filter is given along with a justification based on distance measure concepts. Several authors have previously considered aspects of robust filtering. Kassam and Lim [1] and Kassam, Lim and Cimini [2] considered robust Wiener and matched filters using a band-model for the PSD's. Poor [3, 4] then considered these filters in a more general framework.

The results obtained here are useful in situations where filtering is implemented digitally (e.g. using FFT's). In image processing applications, for example, Wiener filters have often been used [5]. Also, piecewise constant filters may be particularly useful in situations where the filter parameters must be periodically adjusted. This is due to the relatively simple forms of these filters.

II. OPTIMUM PIECEWISE CONSTANT FILTER

As stated in the Introduction, in certain situations it may be advantageous to use a piecewise constant filter. For example, this is particularly helpful when the filter parameters must be periodically adjusted. In this section, we provide a procedure for obtaining the optimum piecewise constant filter.

Considering symmetric PSD's such that $S(w)=S(-w)$ and $N(w)=N(-w)$ with total signal and noise powers, σ_S^2 and σ_N^2, given by

$$\sigma_S^2 = \frac{1}{\pi}\int_0^\infty S(w)dw \qquad (4a)$$

$$\sigma_N^2 = \frac{1}{\pi}\int_0^\infty N(w)dw \qquad (4b)$$

we assume a fixed-form filter that is piecewise constant, i.e.,

$$H(w) = C_j, \quad a_{j-1} \le w < a_j, \quad j=1,2,\ldots m \qquad (5)$$

where the C_j are real, non-negative constants and $\{a_j\}_{j=0}^m$ increasing with $a_o \triangleq 0$ and $a_m \triangleq \Omega$ where Ω is the known bandwidth of the signal. The resulting MSE becomes

$$e(S,N;H) = \frac{1}{\pi}\sum_{j=1}^{m}\int_{a_{j-1}}^{a_j} [S(w)(1-C_j)^2 + N(w)C_j^2]dw \qquad (6)$$

The breakpoints, a_j, and the levels, C_j, then must be chosen to minimize (6).

To minimize the MSE we differentiate it with respect to a_j and C_j. The partial derivative with respect to C_j gives

$$\frac{\partial e}{\partial C_j} = \frac{1}{\pi} \int_{a_{j-1}}^{a_j} [-2(1-C_j)S(w)+2C_j N(w)]dw.$$

Setting this equal to zero and solving for C_j, we obtain

$$C_j = \frac{\int_{a_{j-1}}^{a_j} S(w)dw}{\int_{a_{j-1}}^{a_j} (S(w)+N(w))dw}, \quad 1 \leq j \leq m \qquad (7)$$

which is simply the ratio of the signal power in the interval $[a_{j-1},a_j]$ to the corresponding signal plus noise power. The partial derivative with respect to a_j gives

$$\frac{\partial e}{\partial a_j} = \frac{1}{\pi}[(1-C_j)^2 S(a_j)-(1-C_{j+1})^2 S(a_j)] +$$

$$\frac{1}{\pi}[C_j^2 N(a_j)-C_{j+1}^2 N(a_j)].$$

Setting this equal to zero, we obtain the equation

$$S(a_j)[(2-C_j-C_{j+1})(C_{j+1}-C_j)]=N(a_j)[(C_{j+1}+C_j)(C_{j+1}-C_j)],$$

$$1 \leq j \leq m-1 \qquad (8)$$

(where we are also assuming that $S(w)$ and $N(w)$ are continuous at $w = a_j$, $1 \leq j \leq m$). Equations (7) and (8) are a set of $(2m-1)$ equations which must be satisfied for a piecewise constant filter to possibly be the OPC filter. When considering a filter for which $C_j=C_{j+1}$ for some $j=j_0$, then (8) is always satisfied for $j=j_0$. For other cases, (8) is equivalent to

$$H_o(a_j) = \frac{C_j+C_{j+1}}{2} \qquad (9)$$

(provided we allow $H_o(a_j)$ to have any value when $S(a_j)=N(a_j)=0$).

As stated above, when $C_j=C_{j+1}$ for some j, (8) is always satisfied. However, as the following discussion shows, given the number of allowable breakpoints, this situation is not optimal unless $H_o(w)$ is itself a piecewise constant filter. Consider a piecewise constant filter with $C_j=C_{j+1}$ (for some j, and such that the necessary conditions above are satisfied). For any constant interval $[a,b)$ with level C_{ab} (i.e., either $[a_{j-1},a_{j+1})$ or any other

interval $[a_{i-1}, a_i)$, $i \neq j$), if there exists a subdivision with different levels which results in a lower MSE, then the original filter is not optimal. The MSE associated with the frequency interval $[a,b)$ with filter level C_{ab} is given as

$$e(S,N;H)_{ab} = (1-C_{ab})^2 \sigma_{S_{ab}}^2 + C_{ab}^2 \sigma_{N_{ab}}^2 \qquad (10)$$

where

$$\sigma_{S_{ab}}^2 = \frac{1}{\pi} \int_a^b S(w)dw$$

and

$$\sigma_{N_{ab}}^2 = \frac{1}{\pi} \int_a^b N(w)dw$$

With C_{ab} defined as in (7), the partial MSE of (10) becomes

$$e(S,N;H)_{ab} = \frac{\sigma_{S_{ab}}^2 \sigma_{N_{ab}}^2}{\sigma_{S_{ab}}^2 + \sigma_{N_{ab}}^2} \qquad (11)$$

Now, using a different filter, $H^*(w)$, obtained by dividing the interval $[a,b)$ into $[a,t)$ and $[t,b)$ with levels determined from (7), with

$$\sigma_S^2(t) = \frac{1}{\pi} \int_a^t S(w)dw$$

$$\sigma_N^2(t) = \frac{1}{\pi} \int_a^t N(w)dw$$

[thus $\sigma_S^2(b-t)=\sigma_{S_{ab}}^2 -\sigma_S^2(t)$ and $\sigma_N^2(b-t)=\sigma_{N_{ab}}^2 -\sigma_N^2(t)$]

we obtain the partial MSE

$$e(S,N;H^*)_{ab} = \frac{\sigma_S^2(t)\sigma_N^2(t)}{\sigma_S^2(t)+\sigma_N^2(t)} + \frac{\sigma_S^2(b-t)\sigma_N^2(b-t)}{\sigma_S^2(b-t)+\sigma_N^2(b-t)} \qquad (12)$$

By writing out the MSE expressions and performing some algebraic manipulations we obtain the condition for lowering the MSE. If there exists a point $t\epsilon(a,b)$ such that

$$\frac{\sigma_S^2(t)}{\sigma_N^2(t)} \neq \frac{\sigma_S^2(b-t)}{\sigma_N^2(b-t)} \qquad (13)$$

then the filter with $C_j=C_{j+1}$ is not optimum. On the other hand, if

$$\frac{\sigma_S^2(t)}{\sigma_N^2(t)} = \frac{\sigma_S^2(b-t)}{\sigma_N^2(b-t)} \quad \text{for all } t\epsilon(a,b)$$

then

$$\frac{S(t)}{N(t)} = \frac{\sigma^2_{S_{ab}}}{\sigma^2_{N_{ab}}} \quad \text{for all } t\epsilon(a,b)$$

or, equivalently, $H_o(w)$ is constant over the interval (a,b). Thus, if (13) is not satisfied for all the intervals, $H_o(w)$ is a piecewise constant filter.

For the special case where $C_j = C_{j+1}$ for all j, the filter is a constant for all w and is given as

$$H(w) = \frac{\sigma^2_S}{\sigma^2_S + \sigma^2_N}$$

This filter structure has previously been used (in two-dimensional form) in image processing applications [5] when very little is known about the signal and noise processes.

To obtain the OPC filter parameters, C_j and a_j, we must simultaneously solve (7) and (8) for all j. In all but extremely simple cases these equations are not directly solvable but require an iterative technique for solution. Max [6] developed an algorithm for iteratively solving equations such as these. For the examples which we will consider a simpler searching algorithm for finding the minimum of a function of several variables was applied to the general expression for the MSE (6).

Equations (7) and (8) are only necessary conditions. In general, the sufficient conditions are difficult to obtain. The best answer is that if all the second partial derivatives of the MSE with respect to the levels and the breakpoints exist, then a critical point determined by the necessary conditions is a minimum if the matrix whose ith row and jth column element is

$$\frac{\partial^2 e}{\partial x_i \partial x_j}\Bigg|_{\substack{\text{critical} \\ \text{point}}} , \quad x_j = C_j \text{ or } a_j, \ j=1,2,\ldots m$$

is positive definite [6].

At this point it is interesting to consider the optimum quantization (OQ) of $H_o(w)$. The OQ of $H_o(w)$ is obtained by minimizing

$$\epsilon = \sum_{j=1}^{m} \int_{a_{j-1}}^{a_j} (H_o(w) - C_j)^2 dw \tag{14}$$

Minimizing as before, we obtain the necessary conditions

$$C_j = \frac{\int_{a_{j-1}}^{a_j} H_o(w) dw}{a_j - a_{j-1}} = \begin{array}{l} \text{average value of } H_o(w) \text{ on the} \\ \text{interval } [a_{j-1}, a_j] \end{array}$$

(15a)

and

$$H_o(a_j) = \frac{C_{j+1} + C_j}{2} \tag{15b}$$

which must be solved simultaneously to obtain the OQ of $H_o(w)$. Obviously, the OPC filter and the OQ of $H_o(w)$ are not the same, and, in a linear filtering sense, the OQ filter is not optimum. That is, it is not the best way to obtain a piecewise constant filter given the signal and noise PSD's. If $H_o(w)$ is already a piecewise constant filter, the OPC and OQ filters will coincide.

As an illustration of the design of an OPC filter, we will consider the specific class of signal and noise PSD's possessing the general Butterworth form ($\Omega \to \infty$) i.e.,

$$S(w) = \frac{A}{1 + (\frac{w}{a_S})^{2p}}, \quad P \geq 1 \tag{16a}$$

and

$$N(w) = \frac{B}{1 + (\frac{w}{a_N})^{2n}}, \quad n \geq 1 \tag{16b}$$

where

$$A = \frac{2p\sigma^2_S}{a_S} \sin \frac{\pi}{2p}$$

and

$$B = \frac{2n\sigma^2_N}{a_N} \sin \frac{\pi}{2n}$$

Using these as our nominal densities we determined the OPC filter parameters and calculated the resulting MSE's. Table 1 gives the results for several values of p, n, and a_S. In all cases, we assumed $\sigma^2_S = \sigma^2_N = 1$ and $a_N = 1$. The MSE using the optimum filter has also been tabulated for comparison. These results show the small degradation in performance when using the OPC filter as opposed to the optimum filter, $H_o(w)$. In each case, the optimum filter is very nicely approximated by the OPC filter. That is, the OPC filter provides the benefits obtained from quantizing $H_o(w)$ but with a smaller MSE than if we were to use the OQ filter. One should note that for this particular set of PSD's very little improvement in performance is obtained by increasing m (i.e., finer "quantization").

As an example, we will consider the case p=3, n=1, and $a_S=1$, i.e.,

$$S(w) = \frac{6\sin\frac{\pi}{6}}{1+w^6}$$

and

$$N(w) = \frac{2}{1+w^2}$$

The optimum filter (3) is given as

$$H_o(w) = \cfrac{1}{1+\cfrac{1+w^6}{3\sin\frac{\pi}{6}(1+w^2)}}$$

with the corresponding MSE, $e(S,N;H_o)=0.3934$. The OPC filter ($m=2$) is determined to be

$$H_{OPC} (w) = \begin{cases} 0.615, & |w| < 1.392 \\ 0.080, & |w| \geq 1.392 \end{cases}$$
$$(m=2)$$

with the corresponding MSE, $e(S,N;H_{OPC_{m=2}})=0.4030$ which is a degradation in performance of only 2.4%. When an additional breakpoint is allowed (i.e., $m=3$), the OPC filter is found to be

$$H_{OPC} (w) = \begin{cases} 0.627, & |w| < 1.1844 \\ 0.352, & 1.1844 \leq |w| < 1.6807 \\ 0.039, & |w| \geq 1.6807 \end{cases}$$
$$(m=3)$$

with the corresponding MSE, $e(S,N;H_{OPC_{m=3}})=0.3972$, a degradation from using $H_o(w)$ of less than 1%. As shown in Figure 1, the OPC filter nicely approximates the optimum filter, $H_o(w)$, at the cost of only a small degradation in performance.

III. ROBUSTNESS AND THE p-POINT SPECTRAL CLASS

In order to implement the Wiener filter, the signal and noise PSD's must be known exactly. This is usually an unreasonable assumption and, consequently, small deviations from these assumed, or nominal, PSD's may result in large degradations in performance. In situations such as these we would like to design filters which are less sensitive to deviations from the assumed PSD's.

Essentially, we are interested in obtaining a filter, $H_R(w)$, which will give a non-trivial upper bound for the MSE over the classes of signal and noise PSD's, $\mathcal{S}$ and $\mathcal{N}$. Such a filter is robust in the sense that a certain level of estimation performance is guaranteed for the entire classes, $\mathcal{S}$ and $\mathcal{N}$. Mathematically, the robust Wiener filter, $H_R(w)$, is defined to be the minimax filter such that

$$\min_{\substack{H}} \max_{\substack{S\in\mathcal{S}\\N\in\mathcal{N}}} e(S,N;H) = \max_{\substack{S\in\mathcal{S}\\N\in\mathcal{N}}} e(S,N;H_R) \qquad (17)$$

If $H_R(w)$ is also the optimum filter for a least favorable pair of PSD's, $S_R(w)\in\mathcal{S}$ and $N_R(w)\in\mathcal{N}$, then we have

$$e(S,N;H_R) \leq e(S_R,N_R;H_R) \leq e(S_R,N_R;H) \qquad (18)$$

for any pair $(S,N)\in\mathcal{S}\times\mathcal{N}$ and any linear filter $H(w)$.

The particular classes of PSD's which we will consider are defined as

$$\mathcal{S} \triangleq \{S(\cdot)|\frac{1}{\pi}\int_{a_{j-1}}^{a_j} S(w)dw = P_{S_j}\sigma_S^2, \ j=1,2,3,\ldots m\} \qquad (19a)$$

$$\mathcal{N} \triangleq \{N(\cdot)|\frac{1}{\pi}\int_{a_{j-1}}^{a_j} N(w)dw = P_{N_j}\sigma_N^2, \ j=1,2,3,\ldots m\} \qquad (19b)$$

where $a_o \triangleq 0$, $a_m \triangleq \Omega$, $0 \leq P_{S_j}$, $P_{N_j} \leq 1$ and

$$\sum_{j=1}^m P_{S_j} = \sum_{j=1}^m P_{N_j} = 1$$ and where we are assuming that $S(w)$ and $N(w)$ are symmetric about the origin and continuous at $w=a_j$, $j=1,2,\ldots m$. The breakpoints, a_j, and and the fractional powers, $P_{S_j}\sigma_S^2$ and $P_{N_j}\sigma_N^2$, are assumed known. The mth breakpoint is taken to be infinity unless the bandwidth of the signal is also known. These classes of PSD's are termed p-point spectral classes, and are useful because the fractional power is an easily measured parameter of the process.

Intuitively, it seems that if the only knowledge about the signal and noise PSD's involves the fractional powers, then the least favorable signal and noise PSD's can be anything so long as their ratio is piecewise constant. The previously mentioned image processing application hints at this solution. Also, previous work in robust filtering [1,3] has shown that the least favorable PSD's tend to look as much alike as possible. These intuitions are stated more rigorously in the following theorem.

Theorem 1: For any signal and noise PSD's, $S(w)$ and $N(w)$, which are members of the general p-point classes, $\mathcal{S}$ and $\mathcal{N}$, respectively, the robust filter, $H_R(w)$, satisfying (18) is given by

$$H_R(w) = \frac{S_R(w)}{S_R(w)+N_R(w)} \qquad (20)$$

where the least favorable PSD's, $S_R(w)$ and $N_R(w)$, are such that

$$S_R(w) = k_j N_R(w), \ w\in[a_{j-1},a_j) \qquad (21)$$

where

$$k_j = \frac{\int_{a_{j-1}}^{a_j} S(w)dw}{\int_{a_{j-1}}^{a_j} N(w)dw} = \frac{P_{S_j}\sigma_S^2}{P_{N_j}\sigma_N^2}, \ j=1,2,\ldots,m \qquad (22)$$

Proof: The right-hand side of (18) is true by the definition of $H_R(w)$ as the optimum filter for $S_R(w)$ and $N_R(w)$. Thus, it remains to show that

$$e(S;N;H_R) \leq e(S_R,N_R;H_R) \text{ for all } (S,N)\in\mathcal{S}\times\mathcal{N}.$$

The appropriate MSE's are calculated as

$$e(S_R,N_R;H_R)=\frac{1}{\pi}\sum_{j=1}^m \int_{a_{j-1}}^{a_j}(S_R(w)(\frac{1}{1+k_j})^2+N_R(w)(\frac{k_j}{1+k_j})^2)dw$$

and

$$e(S,N;H_R)=\frac{1}{\pi}\sum_{j=1}^m \int_{a_{j-1}}^{a_j}(S(w)(\frac{1}{1+k_j})^2+N(w)(\frac{k_j}{1+k_j})^2)dw$$

But since $S(w)$ and $S_R(w)$ both belong to $\mathcal{S}$ and $N(w)$ and $N_R(w)$ both belong to $\mathcal{N}$, both quantities reduce to the same result and we obtain

$$e(S,N;H_R)-e(S_R,N_R;H_R) = \sum_{j=1}^{m} \frac{P_{S_j}\sigma_S^2 P_{N_j}\sigma_N^2}{P_{S_j}\sigma_S^2+P_{N_j}\sigma_N^2} \quad (23)$$

and the assertion is proved. Thus, once the robust filter is constructed it gives the same level of performance for all signal and noise PSD's in the general p-point class for which it was defined. Contrast this with the performance of the optimum Wiener filter, $H_0(w)$, which may degrade considerably with small deviations from the assumed PSD's.

It is instructive to consider the statement made earlier that the least favorable PSD's tend to be "closest" in some sense. We will use the concept of distance measures [7,8] to justify this statement. We have the following lemma.

Lemma 1: The least favorable PSD's, $S_R(w)$ and $N_R(w)$, defined as in Theorem 1, minimize the distance measure, $d(S,N)$, over the p-point spectral classes, $\mathcal{S}$ and $\mathcal{N}$, where $d(S,N)$ is defined as†

$$d(S,N) \overset{\Delta}{=} E_N[C(L)] = 2\int_0^\Omega N(w)C(L)dw \quad (24)$$

where $C(\cdot)$ is any real, continuous, convex function on $(0,\infty)$ and $L \overset{\Delta}{=} \frac{S(w)}{N(w)}$.

Proof: Assuming $d(S,N)$ is finite, we must show that the least favorable pair minimizes the distance measure, i.e.,

$$d(S_R,N_R) \leq d(S,N) \text{ for all } (S,N)\epsilon\mathcal{S}x\mathcal{N} \quad (25)$$

Now,

$$\frac{1}{2}[d(S,N)-d(S_R,N_R)] = \sum_{j=1}^{m}[\int_{a_{j-1}}^{a_j} NC(\frac{S}{N})dw - \int_{a_{j-1}}^{a_j} N_R C(\frac{S_R}{N_R})dw]$$

$$= \sum_{j=1}^{m}[\int_{a_{j-1}}^{a_j} N(C(\frac{S}{N})-C(\frac{S_R}{N_R}))dw$$

$$+ \int_{a_{j-1}}^{a_j}(N-N_R)C(\frac{S_R}{N_R})dw] \quad (26)$$

where we are assuming that $0 < S_R(w),N_R(w) < \infty$. But, for convex functions, $C''(w) \geq 0$, so that

$$C(\frac{S}{N}) - C(\frac{S_R}{N_R}) \geq C'(\frac{S_R}{N_R})(\frac{S}{N} - \frac{S_R}{N_R})$$

Equation (26) then becomes (after some manipulation)

$$\frac{1}{2}[d(S,N)-d(S_R,N_R)] \geq \sum_{j=1}^{m}[\int_{a_{j-1}}^{a_j}(N-N_R)\{C(\frac{S_R}{N_R})-C'(\frac{S_R}{N_R})(\frac{S_R}{N_R})\}dw$$

† w dependence has been dropped for convenience.

$$+ \int_{a_{j-1}}^{a_j}(S-S_R)\{C'(\frac{S_R}{N_R})\}dw] \quad (27)$$

However, from Theorem 1, the ratio S_R/N_R is piecewise constant so that (27) becomes

$$\frac{1}{2}[d(S,N)-d(S_R,N_R)] \geq$$

$$\sum_{j=1}^{m}[A_j\int_{a_{j-1}}^{a_j}(N-N_R)dw + B_j\int_{a_{j-1}}^{a_j}(S-S_R)dw]$$

where A_j and B_j are constants. But since $S(w)$ and $S_R(w)$ both belong to $\mathcal{S}$ and $N(w)$ and $N_R(w)$ both belong to $\mathcal{N}$, the right-hand side is zero and we obtain

$$d(S,N) \geq d(S_R,N_R) \text{ for all } (S,N)\epsilon\mathcal{S}x\mathcal{N}$$

the desired result.

Example: A simple (although somewhat contrived) example serves to illustrate the performance invariance of the robust filter. Assume the nominal signal and noise PSD's are given by

$$S_0(w) = \begin{cases} 1/2, & |w| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$N_0(w) = \begin{cases} 1-|w|, & |w| \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

with total signal and noise powers $\sigma_S^2=\sigma_N^2=\frac{1}{2\pi}$. We will assume that these PSD's are really "guesses" based in part on power measurements made at the arbitrarily chosen p-point locations $a_1=1/2$ and $a_2=1$ (the signal bandwidth value which has been assumed known). The optimum filter for these nominal PSD's is obtained as

$$H_0(w) = \frac{1}{3-2|w|}, \ |w| \leq 1$$

with the corresponding minimum MSE, $e(S_0,N_0;H_0) = 0.0717$. The robust filter, $H_R(w)$, designed using only knowledge of the fractional signal and noise powers is obtained as

$$H_R(w) = \begin{cases} 2/5, & 0 \leq |w| < 1/2 \\ 2/3, & 1/2 \leq |w| < 1 \end{cases}$$

with the corresponding MSE, $e(S_0,N_0;H_R)=0.0743$ (a degradation of only ~3.6%). Suppose, now, that the actual signal PSD (let $N(w)=N_0(w)$) is

$$S(w) = \begin{cases} 1, & 0 \leq |w| < 1/4, \ 1/2 \leq |w| < 3/4 \\ 0, & \text{otherwise} \end{cases}$$

Using the previously derived optimum filter, $H_0(w)$, the performance degrades ~13% to a value of $e(S,N_0;H_0)=0.0811$ while the robust filter maintains the same performance level, $e(S,N;H_R)=0.0743$. Obviously, for any deviation in $S_0(w)$ and $N_0(w)$, so long as the resulting PSD's are in the general classes, $\mathcal{S}$ and $\mathcal{N}$, the robust filter will provide the

same level of performance while the optimum filter performance may degrade severely.

Up to this point we have not confronted the question of how to choose the location of the breakpoints. Obviously, if nothing but the fractional powers are known it is quite arbitrary where we place the breakpoints. However, if some additional knowledge about the signal and noise characteristics is available, we can use this knowledge to better choose the breakpoint locations. For example, if "guesses" are available for the signal and noise PSD's we can use the OPC filter as a guide for designing the robust filter.

Several extensions are currently under investigation. The most useful concerns the extension of these results to two-dimensions, which would be applicable in image processing. The robustness concept for the p-point spectral class is also being generalized to include cases where the locations of the signal and noise breakpoints do not coincide, where the location of the breakpoints is uncertain (or, equivalently, where there are errors in the power measurements), and where there is some correlation between the signal and noise.

IV. REFERENCES

1. S.A. Kassam and T.L. Lim, "Robust Wiener Filters," J. Franklin Institute, October/November, 1977, pp. 171-185.
2. S.A. Kassam, T.L. Lim, and L.J. Cimini, "Two-dimensional Signal Filters Under Modeling Uncertainties," Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing, April, 1979, pp. 638-641.
3. H.V. Poor, "Some Results in Robust Wiener Filtering," Proc. IEEE Conf. on Decision and Control, January, 1979.
4. H.V. Poor, "A General Approach to Robust Matched Filtering," Proc. Midwest Symposium on Circuits and Systems, June, 1979.
5. R.C. Gonzalez and P. Wintz, Digital Image Processing, Addison-Wesley Publishing Company,Inc. Reading, MA, 1977.
6. J. Max, "Quantizing for Minimum Distortion," IEEE Trans. on Information Theory, Vol. 6, March, 1960, pp. 7-11.
7. S.A. Kassam, "An Extension of Huber's Results on Robust Tests," Proc. Midwest Symposium on Circuits and Systems, June, 1979.
8. S.M. Ali and S.D. Silvey, "A General Class of Coefficients of Divergence of One Distribution from Another," J. Royal Stat. Soc., Vol. 28, 1966, pp. 131-142.
9. L.J. Cimini, "Minimax Estimation Filters," M.S. Thesis, University of Pennsylvania, August, 1979.

TABLE 1. OPC filter parameters and corresponding MSE's for Butterworth PSD's (MSE's using $H_o(w)$ is given for comparison) [9]

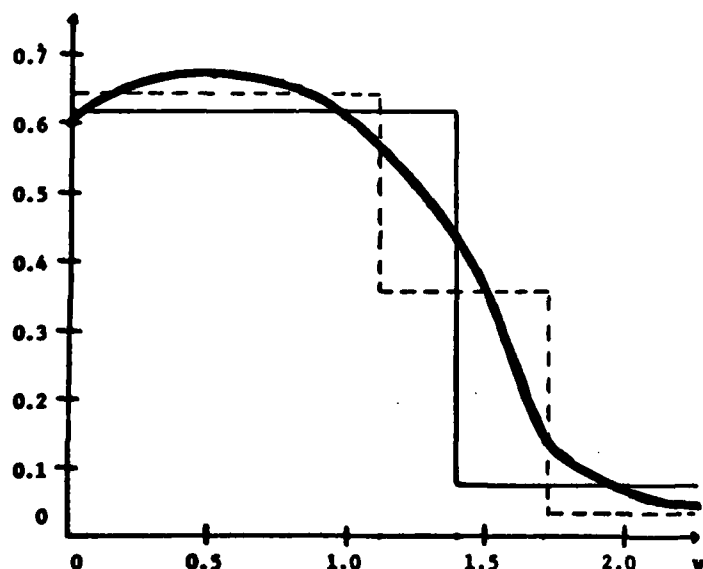| p	n	a_S	m	a_1	c_1	MSE	MMSE($H_o(w)$)
3	1	1	2	1.392	0.615 0.080	0.4030	0.3934
3	1	1	3	1.184 1.681	0.627 0.352 0.039	0.3972	0.3934
3	1	.5	2	0.670	0.718 0.063	0.3085	0.2939
9	1	1	2	1.103	0.651 0.022	0.3557	0.3499
9	1	1	3	1.046 1.165	0.655 0.351 0.009	0.3531	0.3499
9	1	1	4	1.038 1.153 1.810	0.655 0.396 0.036 0.001	0.3531	0.3499
9	1	2	2	2.277	0.574 0.022	0.4291	0.4084
4	2	1	2	1.243	0.529 0.177	0.4821	0.4775
2	4	1	2	1.249	0.471 0.830	0.4821	0.4775

Figure 1. OPC filter for Butterworth PSD's, $p=3, n=1, a_S=1$.

$H_o(w)$

$H_R(w)$ (m=2)

$H_R(w)$ (m=3)